Homework 2

Problem 1. Let $Y \in \{1,2,3,4,5,6\}$ be the face number showing when a die is rolled. Define X as

$$X = \begin{cases} Y & \text{if } Y \text{ is even,} \\ 0 & \text{if } Y \text{ is odd.} \end{cases}$$

Find the best linear predictor $\mathcal{P}(Y|X)$ and the conditional expectation $\mathbb{E}(Y|X)$. Calculate $\mathbb{E}\left[(Y - \mathcal{P}(Y|X))^2\right]$ and $\mathbb{E}\left[(Y - \mathbb{E}(Y|X))^2\right]$.

Problem 2. Suppose that

$$\begin{aligned} Y &= \boldsymbol{X}'\boldsymbol{\beta} + e \\ \mathrm{E}(e|\boldsymbol{X}) &= 0 \\ \mathrm{E}\left(e^2|\boldsymbol{X}\right) &= \sigma^2(\boldsymbol{X}). \end{aligned}$$

Consider two approximations to the conditional variance $\sigma^2(\boldsymbol{X})$:

$$\boldsymbol{\gamma}_1 \text{ minimizes } \mathrm{E}\left(\sigma^2(\boldsymbol{X}) - \boldsymbol{X}'\boldsymbol{\gamma}\right)^2$$

and

$$\boldsymbol{\gamma}_2 \text{ minimizes } \mathrm{E}\left(e^2 - \boldsymbol{X}'\boldsymbol{\gamma}\right)^2.$$

Show: $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_2$.

Problem 3. The Mean Trimmed Squared Error (MTSE) is defined by

$$T(\boldsymbol{\theta}) = \mathrm{E}\left(\left(Y - \boldsymbol{X}'\boldsymbol{\theta}\right)^2 \tau(\boldsymbol{X})\right),$$

where $\tau(\boldsymbol{X})$ is a known, scalar-valued, non-negative, bounded, function.

1. Give an explicit formula for the value of $\boldsymbol{\theta}$ which minimizes $T(\boldsymbol{\theta})$.
2. Define $e = Y - \boldsymbol{X}'\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ is the minimizer defined above. Show: $\mathrm{E}(\boldsymbol{X}\tau(\boldsymbol{X})e) = 0$.
3. Under what condition (other than $\tau(\boldsymbol{X}) = 1$) will this minimizer equal the Best Linear Predictor?

Problem 4. The conditional PDF of Y given X is

$$f_{Y|X}(y|x) = \frac{y + x}{1/2 + x},$$

for $0 < y < 1$. Find $\mathrm{E}[Y|X = x]$.

Problem 5. For any given two random variables X and Y, we define

$$\mathrm{Var}[Y \mid X] = \mathrm{E}\left[(Y - \mathrm{E}[Y \mid X])^2 \mid X\right].$$

Suppose that $\mathrm{E}[Y \mid X] = 1/4$ and $\mathrm{E}\left[Y^2 \mid X\right] = 1/8$. Then show that for any function g, $\mathrm{Var}[Y \mid g(X)] = 1/16$. Use the following facts: for any function g, $\mathrm{E}[\mathrm{E}[Y \mid g(X)] \mid X] = \mathrm{E}[Y \mid g(X)]$ and $\mathrm{E}[\mathrm{E}[Y \mid X] \mid g(X)] = \mathrm{E}[Y \mid g(X)]$.

Problem 6. Let $\boldsymbol{X}$ be the matrix collecting all the n observations on the k regressors:

$$\boldsymbol{X} = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,k} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n,1} & X_{n,k} & \cdots & X_{n,k} \end{bmatrix}_{n \times k}.$$

Assume $n > k$ and $\boldsymbol{X}$ is of full rank. Let $\boldsymbol{A}$ be a $k \times k$ *singular* matrix. Show that the columns of $\boldsymbol{XA}$ are linearly dependent and $\mathcal{S}(\boldsymbol{XA}) \subset \mathcal{S}(\boldsymbol{X})$, where

$$\mathcal{S}(\boldsymbol{X}) = \left\{ \boldsymbol{z} \in \mathbb{R}^n : \boldsymbol{z} = \boldsymbol{Xb},\, \boldsymbol{b} = (b_1, b_2, \ldots, b_k)' \in \mathbb{R}^k \right\}.$$

Problem 7. Partition the matrix of regressors $\boldsymbol{X}$ as follows:

$$\boldsymbol{X} = [\boldsymbol{X}_1 \ \ \boldsymbol{X}_2].$$

Denote $\boldsymbol{P}_1 = \boldsymbol{X}_1 \left(\boldsymbol{X}_1'\boldsymbol{X}_1\right)^{-1} \boldsymbol{X}_1'$ and $\boldsymbol{P}_{\boldsymbol{X}} = \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'$. $\boldsymbol{M}_1$ and $\boldsymbol{M}_{\boldsymbol{X}}$ are defined analogously: $\boldsymbol{M}_1 = \boldsymbol{I}_n - \boldsymbol{P}_1$ and $\boldsymbol{M}_{\boldsymbol{X}} = \boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}}$. Prove:

$$\boldsymbol{P}_1\boldsymbol{P}_{\boldsymbol{X}} = \boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{P}_1 = \boldsymbol{P}_1 \tag{1}$$

and

$$\boldsymbol{M}_1\boldsymbol{M}_{\boldsymbol{X}} = \boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{M}_1 = \boldsymbol{M}_{\boldsymbol{X}}. \tag{2}$$

Problem 8. Use (1) to show that $\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_1$ is symmetric and idempotent. Show further that $\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_1 = \boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}$ by showing that for any $\boldsymbol{z} \in \mathcal{S}(\boldsymbol{M}_1\boldsymbol{X}_2)$, $(\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_1)\boldsymbol{z} = \boldsymbol{z}$ and for any $\boldsymbol{y} \in \mathcal{S}^{\perp}(\boldsymbol{M}_1\boldsymbol{X}_2)$, $(\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_1)\boldsymbol{y} = \boldsymbol{0}$, where

$$\mathcal{S}^{\perp}(\boldsymbol{M}_1\boldsymbol{X}_2) = \{\boldsymbol{z} \in \mathbb{R}^n : \boldsymbol{z}'\boldsymbol{M}_1\boldsymbol{X}_2 = \boldsymbol{0}\}.$$

Problem 9. In this question, use the hints to show "R^2 increases by adding more regressors". Suppose we have n observations on regressors $(Z_1, ..., Z_k)$ and $(W_1, ..., W_m)$ and dependent variable Y. Let $\boldsymbol{Z}$ be the $n \times k$ matrix collecting the observations on $(Z_1, ..., Z_k)$ and let $\boldsymbol{W}$ be the $n \times m$ matrix collecting the observations on $(W_1, ..., W_m)$. Let $\boldsymbol{X} = [\boldsymbol{Z} \ \ \boldsymbol{W}]$. Assume that $\boldsymbol{Z}$ contains a column of ones so that $R^2 = 1 - RSS/TSS$ in both regressions.

Let

$$\boldsymbol{P}_{\boldsymbol{X}} = \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}' \text{ projection matrix corresponding to the full regression,}$$
$$\boldsymbol{P}_{\boldsymbol{Z}} = \boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}' \text{ projection matrix corresponding to the regression without } \boldsymbol{W}.$$

Define also

$$\boldsymbol{M}_{\boldsymbol{X}} = \boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{X}},$$
$$\boldsymbol{M}_{\boldsymbol{Z}} = \boldsymbol{I}_n - \boldsymbol{P}_{\boldsymbol{Z}}.$$

Define

$$\widehat{\boldsymbol{e}}_{\boldsymbol{X}} = \boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{Y},$$
$$\widehat{\boldsymbol{e}}_{\boldsymbol{Z}} = \boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{Y}.$$

Show: $\widehat{\boldsymbol{e}}_{\boldsymbol{X}}'\widehat{\boldsymbol{e}}_{\boldsymbol{Z}} = \widehat{\boldsymbol{e}}_{\boldsymbol{X}}'\widehat{\boldsymbol{e}}_{\boldsymbol{X}}$ and therefore

$$0 \leq (\widehat{\boldsymbol{e}}_{\boldsymbol{X}} - \widehat{\boldsymbol{e}}_{\boldsymbol{Z}})'(\widehat{\boldsymbol{e}}_{\boldsymbol{X}} - \widehat{\boldsymbol{e}}_{\boldsymbol{Z}}) = \widehat{\boldsymbol{e}}_{\boldsymbol{X}}'\widehat{\boldsymbol{e}}_{\boldsymbol{X}} - \widehat{\boldsymbol{e}}_{\boldsymbol{Z}}'\widehat{\boldsymbol{e}}_{\boldsymbol{Z}}.$$

Hint: use (1) and (2). How can you argue that now we conclude that "R^2 increases by adding more regressors"?

Problem 10. Let $\boldsymbol{X}$ be an $n \times k$ matrix $(n > k)$ of full column rank. Partition $\boldsymbol{X}$ as $\boldsymbol{X} = [\ \boldsymbol{X}_1 \quad \boldsymbol{X}_2\ ]$, where $\boldsymbol{X}_1$ is $n \times k_1$ and $\boldsymbol{X}_2$ is $n \times k_2$, $k_1 + k_2 = k$.

1. Show that $\boldsymbol{X}_2$ has full column rank and therefore $(\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}$ exists.
2. Define $\boldsymbol{M}_2 = \boldsymbol{I}_n - \boldsymbol{X}_2\left(\boldsymbol{X}_2'\boldsymbol{X}_2\right)^{-1}\boldsymbol{X}_2'$ and $\widetilde{\boldsymbol{X}}_1 = \boldsymbol{M}_2\boldsymbol{X}_1$. Show that $\widetilde{\boldsymbol{X}}_1$ has full column rank and therefore $\left(\widetilde{\boldsymbol{X}}_1'\widetilde{\boldsymbol{X}}_1\right)^{-1} = \left(\boldsymbol{X}_1'\boldsymbol{M}_2\boldsymbol{X}_1\right)^{-1}$ exists.